(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number : **0 526 347 A3**

(12) EUROPEAN PATENT APPLICATION

(21) Application number : **92402208.0**

(51) Int. Cl.^{5} : **G10L 5/06**

(22) Date of filing : **31.07.92**

(30) Priority : **01.08.91 JP 193261/91**

(43) Date of publication of application :
03.02.93 Bulletin 93/05

(84) Designated Contracting States :
DE FR GB

(88) Date of deferred publication of search report :
11.08.93 Bulletin 93/32

(71) Applicant : **FUJITSU LIMITED**
1015, Kamikodanaka Nakahara-ku
Kawasaki-shi Kanagawa 211 (JP)

(72) Inventor : **Hamasaki, Ryosuke, c/o FUJITSU LIMITED**
1015, Kamikodanaka, Nakahara-ku
Kawasaki-shi, Kanagawa 211 (JP)
Inventor : **Loken-Kim, Kyung-Ho, c/o FUJITSU LIMITED**
1015, Kamikodanaka, Nakahara-ku
Kawasaki-shi, Kanagawa 211 (JP)

(74) Representative : **Joly, Jean-Jacques et al**
Cabinet Beau de Loménie 158, rue de l'Université
F-75340 Paris Cédex 07 (FR)

(54) A number-of-recognition candidates determining system in a speech recognizing device.

(57) A number-of-recognition candidates determining system in a speech recognition device, for calculating the similarity between a plurality of preliminarily registered speeches and an inputted speech and outputting as recognition candidates a predetermined number of said registered speeches at higher order in similarity, comprises an SN ratio calculator (15) for calculating an SN ratio of an inputted speech to the environmental noise, and a number-of-recognition candidates controller (16) for controlling the number of the recognition candidates according to the value of the SN ratio calculated by the SN ratio calculator. The number of recognition candidates is varied according to the value of the SN ratio.

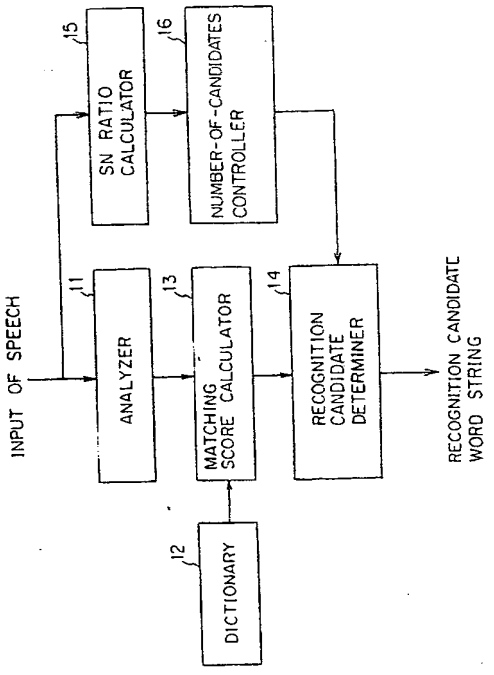

Fig. 3

EP 0 526 347 A3

BNSDOCID: <EP_____0526347A3_I_>

European Patent Office

EUROPEAN SEARCH REPORT

Application Number

EP 92 40 2208

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. Cl.5)
Y	US-A-4 860 358 (L.R. RABINER) * Column 5, line 62 - column 6, line 1; column 7, lines 55-67 *	1	G 10 L 5/06
P,Y	EP-A-0 487 307 (NEC CORP.) * Claim 3 *	1	
A	US-A-4 720 802 (J.N. DAMOULAKIS et al.) * Column 27, line 66 - column 28, line 2 *	1	
A	DE-A-3 819 178 (RICOH CO.) * Column 7, lines 10-17 *	1	

TECHNICAL FIELDS SEARCHED (Int. Cl.5)

G 10 L 5/06

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	17-05-1993	ARMSPACH J F A

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P0401)

BNSDOCID: <EP_____0526347A3_I_>

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number : **0 526 347 A2**

(12) EUROPEAN PATENT APPLICATION

(21) Application number : **92402208.0**

(51) Int. Cl.[5] : **G10L 5/06**

(22) Date of filing : **31.07.92**

(30) Priority : **01.08.91 JP 193261/91**

(43) Date of publication of application :
03.02.93 Bulletin 93/05

(84) Designated Contracting States :
DE FR GB

(71) Applicant : **FUJITSU LIMITED**
1015, Kamikodanaka Nakahara-ku
Kawasaki-shi Kanagawa 211 (JP)

(72) Inventor : **Hamasaki, Ryosuke, c/o FUJITSU LIMITED**
1015, Kamikodanaka, Nakahara-ku
Kawasaki-shi, Kanagawa 211 (JP)
Inventor : **Loken-Kim, Kyung-Ho, c/o FUJITSU LIMITED**
1015, Kamikodanaka, Nakahara-ku
Kawasaki-shi, Kanagawa 211 (JP)

(74) Representative : **Joly, Jean-Jacques et al**
CABINET BEAU DE LOMENIE 55, rue d'Amsterdam
F-75008 Paris (FR)

(54) A number-of-recognition candidates determining system in a speech recognizing device.

(57) A number-of-recognition candidates determining system in a speech recognition device, for calculating the similarity between a plurality of preliminarily registered speeches and an inputted speech and outputting as recognition candidates a predetermined number of said registered speeches at higher order in similarity, comprises an SN ratio calculator (15) for calculating an SN ratio of an inputted speech to the environmental noise, and a number-of-recognition candidates controller (16) for controlling the number of the recognition candidates according to the value of the SN ratio calculated by the SN ratio calculator. The number of recognition candidates is varied according to the value of the SN ratio.

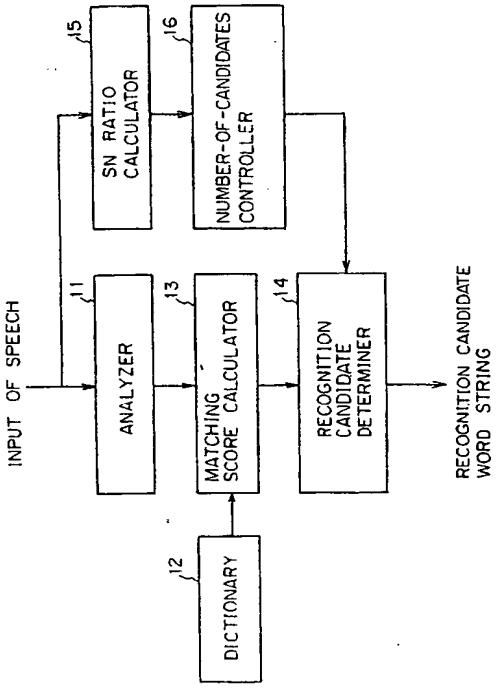

Fig. 3

EP 0 526 347 A2

BNSDOCID: <EP_____0526347A2_I_>

Background of the Invention

The present invention relates to a speech recognition device for recognizing a vocally inputted instruction as data and more particularly to a number-of-recognition candidates determing system in the speech recognizing device for outputting a plurality of recognition candidates by calculating a similarity between a plurality of a preliminarily registered speeches and an inputted speech.

Recently, technologies for recognizing speech using a computer and for vocally inputting instructions and data have been popularly utilized in many fields of industry. The technologies include voice controlled word processors and automatically operated machines, etc.

Figure 1 is a general view of a speech recognizer, and a brief explanation about its operation is given below by referring to Figure 1.

Inputted speech data are analyzed by an analyzer 1, and the analysis result and the feature vectors of each word stored in a dictionary 2 are matched by a matching score calculator 3 to obtain a matching score for "similarity". "distance" is also used instead of "similarity". A small distance means a high similarity. A recognition candidate determiner 4 outputs a recognition candidate word string in the order from the word having the highest similarity according to a predetermined number of word candidates. Assuming that the number of candidates is "10", ten highest similarity words are outputted as recognition candidates.

The recognition rate of the above described speech recognition device greatly depends on the value of an SN ratio, that is, a ratio of an inputted speech to an environmental noise. For example, a speech recognition device inputs different SN ratios depending on its location and the time because an environmental noise changes with time of day.

Figure 2 is a graph indicating the change in the recognition rate of each SN ratio depending on the number of candidates to be recognized. In Figure 2, the recognition rate at the SN ratio of 50dB is represented by the solid line, that at the SN ratio of 30dB is represented by a broken line, and that at the SN ratio of 10dB is represented by a dashed line. In Figure 2, if the number of candidates to be recognized is determined to be "10", the SN ratio of 50dB yields the recognition rate of 95%, that of 30dB yields the rate of 88%, and that of 10dB yields the rate of 82%. Thus, the higher SN ratio indicating a lower noise yield a higher recognition rate.

As described above, an SN ratio of a great value allows a high percentage of recognition, and vice versa. That is, a word having the highest similarity at an SN ratio of a small value can be ranked at a lower priority, for example, the fifteenth, if an SN ratio indicates a small value. In this case, the word is excluded from the candidates if the number of candidates to be recognized is set to 10.

This is caused by the fixed number of word candidates, for example, 10. In this case, ten highest order words are determined to be word candidates and the following words are determined to be excluded from the candidates. Therefore, a word to be reasonably defined as a candidate is often determined to be out of the range of candidates. Thus, conventional speech recognition device yield different recognition rates greatly depending on an SN ratio which is changeable according to the state of its environment.

Summary of the Invention

The present invention aims at calculating an SN ratio changeable by the state of the environment and improving a recognition rate by changing the number of word candidates depending on the change of an SN ratio.

The present invention is used in a speech recognition device for recognizing a speech, namely, a vocally inputted instruction and data, and utilizes a number-of-recognition candidates determining system in the speech recognition device for outputting a plurality of candidates to be recognized by calculating the similarity between a plurality of stored speeches and an inputted speech, and selecting as recognition candidates a plurality of words having higher similarity.

The present invention comprises an SN ratio calculator for calculating an SN ratio, that is, a ratio of a sound level to the environmental noise, and a number-of-candidates controller for controlling the number of word candidates to be recognized.

It further comprises an analyzer, a score calculator, a recognition candidate determiner, and a dictionary. The analyzer analyzes inputted speech, and the analysis result and the feature vectors of a word stored in the dictionary are matched by the score calculator to calculate a score for the "similarity".

A number-of-candidates control signal is outputted from the number-of-candidates controller to the recognition candidate determiner according to an SN ratio obtained by the SN ratio calculator. The recognition candidate determiner outputs a predetermined number of words having a higher priority according to their similarity as a recognition candidate word string. The number of word candidates to be recognized can be controlled corresponding to a given SN ratio. That is, at an SN ratio of a small value, the number of word candidates to be recognized is increased.

Such control prevents the problem that a word, which is ranked at a high priority at an SN ratio of a large value, may be assigned a low priority and is excluded from a candidate group. Thus, a word to be reasonably defined as a candidate can be prevented from being excluded from a candidate group, and can be processed as a word candidate, thereby improving a recognition rate.

BNSDOCID: <EP_____0526347A2_I_>

Brief Description of the Drawings

One of skill in the art can easily understand additional features and objects of this invention from the description of the preferred embodiments and some of the attached drawings. In the drawings:

Figure 1 shows the configuration of a conventional number-of-candidates determining system ;

Figure 2 shows the change in a recognition rate due to the change in an SN ratio when the number of candidates is fixed to a predetermined value in the conventional system;

Figure 3 is a block diagram for explaining the principle of the present invention;

Figure 4 shows the configuration for explaining the first embodiment of the present invention;

Figure 5 shows the configuration of the feature vectors extractor;

Figure 6 shows the configuration of the SN ratio calculator;

Figure 7 shows the relationship between each SN ratio for the recognition rate of 90% and the number of candidates to be recognized;

Figure 8 is a graph for explaining the necessary recognition candidate determiner according to the first embodiment;

Figure 9 shows the configuration for explaining the second embodiment;

Figure 10 is a graph for explaining the necessary recognition candidate determiner according to the second embodiment;

Figure 11 shows the configuration for explaining the third embodiment; and

Figure 12 is a graph for explaining the necessary recognition candidate determiner according to the third embodiment.

Description of the Preferred Embodiment

Figure 3 is a block diagram for explaining the principle of the present invention. As shown in Figure 3, the present invention basically comprises an analyzer or analyzing section 11 for analyzing in inputted speech and extracting its feature vectors i.e. feature vectors, a dictionary 12 for preliminary storage of a plurality of words, a matching score calculator or calculating section 13 for collating an analysis result of the analyzer 11 with the feature vectors i.e. feature vectors of each word stored in the dictionary 12 and calculating a score for "similarity", a recognition candidate determiner or determining section 14 for determining and outputting the predetermined number of highest similarity words, an SN ratio calculator or calculating section 15 for calculating an SN ratio of inputted speech to environmental noise, and a number-of-candidates controller or controlling section 16 for controlling the number of candidates determined by the recognition candidate determiner 14 according to a value of an SN ratio calculated by the SN ratio calculator 15.

With the above described configuration, an inputted speech signal is analyzed by the analyzer 11, the analysis result and the feature vectors of each word stored in the dictionary 12 are matched by the matching score calculator 13, and a score is obtained for "similarity".

The SN ratio calculator 15 calculates an SN ratio of an inputted speech signal to environmental noise, and sends the calculation result to the number-of-candidates controller 16. The number-of-candidates controller 16 outputs to the recognition candidate determiner a number-of-candidates control signal assigned according to the value of a calculated SN ratio .

The recognition candidate determiner 14 determines and outputs the number of candidates according to the above described control signal. That is, after the calculation of an SN ratio, the number of word candidates can be varied depending on the value of an SN ratio. For example, a large value of an SN ratio sets the number of candidates to a predetermined value, 10 for example, and a small value of an SN ratio extends the number of candidates, to 20 words for example.

Such control prevents the problem that a word having the first or second priority in similarity at a high SN ratio may be assigned a low priority, for example the 15th, and excluded from the candidates at a low SN ratio due to an undesirable environmental change. That is, a word to be reasonably defined as a candidate can be processed as a word candidate without being excluded from the candidates by extending the range of the number of word candidates to 20 words, for example, thereby improving a recognition rate.

An embodiment of the present invention is explained below by referring to the drawings. Figure 4 shows the configuration of the first embodiment of the present invention. In Figure 4, the configuration is generally the same as that shown in Figure 3 for explaining the principle of the present invention, but further comprises an AD converter 21 for converting an analog signal to a digital signal, a feature vectors extractor or extracting section 22 for analyzing an AD-converted speech signal at predetermined time intervals and outputting feature vectors used in matching them with those of each word stored in the dictionary 12, a score sorter 23 for sorting word candidates relative to the similarity according to the scores calculated by the matching score calculator 13, a number-of-recognition candidates storing table 24 for storing the number of word candidates for obtaining a predetermined recognition rate at an SN ratio calculated by the SN ratio calculator 15, and an interpolating equation deriver 25 for deriving an interpolating equation for obtaining the number of word candidates stored in the number-of-recognition candidates storing table 24.

BNSDOCID: <EP_____0526347A2_I_>

The AD converter 21 and the feature vectors extractor 22 correspond to the analyzer 11 in Figure 3. The number-of-recognition candidates storing table 24 and the interpolating equation deriver 25 correspond to the number-of-candidates controller 16 in Figure 3.

A technique of analyzing an inputted speech at time intervals by the analyzer 11 is well-known, and can be realized by configuring the feature vectors extractor 22 as shown in Figure 5. The technique shown in Figure 5 refers to an analysis using a digital filter bank. That is, a plurality of digital bandpass filters BPF1 - BPF n each having a different center frequency are provided in the speech frequency band to be analyzed. Then, a rectifying and smoothing component in the output of each bandpass filter is sampled every few milliseconds or every tens of milliseconds, and expressed in logarithm as a feature vectors parameter. The result is outputted to the score calculator 13 to be compared in similarity with words stored in the dictionary 12. An input speech can also be analyzed by using an analog bandpass filter bank, LPC (Linear Prediction) analysis, cepstrum analysis or FFT (Fast Fourier Transform) analysis.

As shown in Figure 6, the SN ratio calculator 15 comprises a speech segment detector or detecting section 151, a speech segment average power calculator or calculating section 152, a noise range average power calculator or calculating section 153, and a logarithmic power ratio calculator or calculating section 154. On receiving at the speech segment detector 151 an inputted speech signal which is converted to the digital speech signal by the AD converter 21, the SN ratio calculator 15 calculates at the speech segment average power calculator 152 a speech segment average power according to a signal from the speech segment detector 151, and calculates a noise range average power at the noise range average power calculator 153. Then, it calculates at the logarithmic power ratio calculator 154 a logarithmic power ratio from the speech segment average power and the noise range average power, and then outputs the result as an SN ratio.

Figure 7 is a graph showing the change in the recognition rate for the number of word candidates at SN ratios of 50dB, 30dB, and 10dB. As shown in Figure 7 and as described above, the smaller value an SN ratio indicates, the lower the recognition rate becomes if the number of word candidates is 10. For example, 50dB (solid line in Figure 7) yields the recognition rate of 95%, 30dB (broken line in Figure 7) yields that of 88%, and 10dB (dashed line in Figure 7) yields that of 82%. Whatever value an SN ratio may indicate, the number of word candidates must be 6 at 50 dB, 20 at 30dB, and 30 at 10 dB to obtain the recognition rate over 90% according to Figure 7. Thus, the threshold of an SN ratio is set to 50dB, 30dB, and 10dB, and the necessary number of word candidates are set to 6, 20, and 30 respectively. The recognition rate represents a possibility that the right candidates is included in a group of the word candidates. The above described values of the threshold of an SN ratio are interpolated by the interpolating equation provided by the interpolating equation deriver 25, and the appropriate number of word candidates corresponding to the SN ratio of an inputted speech can be determined by the number-of-recognition candidates determiner 26 according to the above described interpolating equation. Then, the determined number of words is given to the recognition candidate determiner 14 to output the determined number of words as a recognition candidate word string starting from the first priority in similarity.

The total operation performed with the configuration above is explained below. First, inputted speech is converted to the digital speech signal by the AD converter 21, and analyzed at appropriate time intervals by the feature vectors extractor 22. Then, the score calculator 13 matches the feature vectors of a word stored in the dictionary 12 with the above described analysis result, and the score sorter 23 sorts word candidates in the order of high similarity and transmits them to the recognition word candidate determiner 14.

Assume that an SN ratio of an inputted speech signal calculated by the SN ratio calculator 15 is 20dB.

To obtain the recognition rate over 90%, the threshold of an SN ratio is set to 50dB, 30dB, and 10dB, and the necessary number of word candidates is set to 6, 20, and 30 respectively. The first embodiment, as shown in Figure 8, is an example of interpolating the values of the SN ratios of 50dB, 30dB, and 10dB with a predetermined number of word candidates. Figure 8 shows the number of word candidates for the SN ratios of 50dB, 30dB, and 10dB to obtain the recognition rate over 90%. When the necessary numbers of recognition candidates for the SN ratios of 50dB, 30dB, and 10dB are calculated, the necessary number of recognition candidates between the specified SN ratios is determined to be constant as the number at the lower ratio, thereby generating a graph indicating step changes in the number of word candidates.

For example, if the SN ratio of inputted speech is obtained as 20dB by the SN ratio calculator 15, it is between 10dB and 30dB and the number of word candidates is determined to be assigned 30. Thus, when the necessary number of candidates at the SN ratio of 20dB is determined, the recognition candidate determiner 14 outputs the determined number of words as a recognition candidate word string.

To realize this, the interpolation equation deriver 25 is operated as follows in the first embodiment. The interpolating equation deriver 25 comprises an SN ratio range determiner 251 for determining to what ranges an SN ratio calculated by the SN ratio calcu-

BNSDOCID: <EP_____0526347A2_I_>

lator 15 belongs in the ranges of the above described thresholds 50dB, 30dB, and 10dB. For example, as described above, if an SN ratio calculated by the SN ratio calculator 15 is 20dB, the SN ratio range determiner 251 determines the SN range to be between 30dB and 10dB. When the range of the SN ratio is determined, the corresponding number of recognition candidates (30 in this case) is read, and the data are transmitted to the recognition candidate determiner 14. The interpolating equation in the first embodiment is represented as y = e (where y indicates the number of recognition candidates; e indicates the range of the SN ratio). Therefore, if the SN ratio calculated by the SN ratio calculator 15 is x_1, e (the range of the SN ratio) is determined to be the range of the SN ratio x_1, and the number of recognition candidates y_1 can be determined immediately.

Thus, a recognition rate over a predetermined value (90% in this case) can be obtained at each of the SN ratios of inputted speech data by defining the number of word candidates at an SN ratio between predetermined points of SN ratios as indicating a fixed number of word candidates and changing in the form of steps as shown in Figure 8. In this case, the number of word candidates stored in the number-of-candidates storing table 24 increases in the form of steps, and the predetermined value points of SN ratios, for example 6, 20, and 30 are also change points. Therefore, the number of word candidates can be determined only by comparing the value of an SN ratio of an inputted speech with a predetermined threshold, thereby removing an operation of obtaining a complicated interpolating equation by the interpolating equation deriver 25. Furthermore, since it is not necessary to determine the number of word candidates corresponding to each of the SN ratios for inputted speech, the system does not require a large memory capacity at all.

Next, the second embodiment is explained below by referring to Figures 9 and 10. As shown in Figure 10, each of the points indicating the numbers of necessary word candidates 6, 20, and 30 at SN ratios of an inputted speech 50dB, 30dB, and 10dB respectively are connected by lines to make a polygonal line graph and interpolate the values by a non-linear function.

In this case, as described above by referring to the data of the number-of-recognition candidates storing table 24, the interpolating equation 25 derives an interpolating equation for interpolating values obtained by connecting predetermined points of the number of word candidates 6, 20, and 30 corresponding to the SN ratios 50dB, 30dB, and 10dB respectively to obtain a graph represented by a line polygonal at each of the points as shown in Figure 10.

To realize this, in the second embodiment, the SN ratio range determiner 251 in the interpolating deriver 25 shown in Figure 9, determines where an SN ratio obtained by the SN ratio calculator 15 is located in the threshold range 50dB-30dB-10dB of an SN ratio, and it obtains an interpolating equation to interpolate values as indicated by a polygonal line connecting each of the points. The system also comprises a number-of-recognition candidates determiner 26 for calculating the number of recognition candidates by substituting the value of an SN ratio for the interpolating equation. The interpolating equation in the second embodiment can be represented as follows.

$$y - y_p = \frac{y_{p+1} - y_p}{x_{p+1} - x_p}(x - x_p) \quad (1)$$

where

x_p: a threshold of an SN ratio

x_{p+1}: a threshold of an SN ratio of the value following x_p

y_p: the number of recognition candidates for x_p

y_{p+1}: the number of recognition candidates for x_{p+1}

x: an SN ratio obtained by the SN ratio calculator 15

y: the number of recognition candidates to be obtained

The necessary number of recognition candidates at the SN ratio of x_1 can be calculated by substituting for x in the above equation (1) the SN ratio x_1 calculated by the SN ratio calculator 15. If the necessary number of recognition candidates is y_1, the above described equation (1) can be represented as follows.

$$y_1 - y_p = \frac{y_{p+1} - y_p}{x_{p+1} - x_p}(x_1 - x_p) \quad (2)$$

Equation (2) calculates the number of recognition candidates determined by the number-of-recognition candidates determiner 26 for determining the necessary number of recognition candidates at a specific SN ratio. That is, if an SN ratio is determined to be 20dB, the "20dB" is substituted for x_1 in equation (2). Then, the necessary number of recognition candidates y_1 at the SN ratio can be calculated. In this case, 20dB is located in the SN ratio range between 30dB and 10dB. Therefore, 30dB is substituted for x_{p+1}, 10dB is substituted for x_p, the number of recognition candidates 30 is substituted for y_{p+1}, and 20 is substituted for y_p in equation (2). This yields the result y_1 = 25. Accordingly, when an SN ratio is 20dB, the necessary number of recognition candidates is 25. If the value of y_1 is provided for the recognition candidates determiner 14, the recognition candidate determiner 14 outputs 25 higher order words among the word candidates sorted by the sorter 23 as a recognition candidate word string. If the 25 higher order words are determined to be word candidates, the recognition rate of 90% can be guaranteed.

Thus, since the necessary number of word candidates is varied in proportion to the change in the value of an SN ratio of inputted speech in the second embodiment, the number of word candidates corre-

BNSDOCID: <EP_____0526347A2_I_>

sponding to the value of the SN ratio can be obtained, thereby improving the efficiency in a recognizing operation.

Next, the third embodiment is explained by referring to Figures 11 and 12. In the third embodiment, each of the points indicating the necessary numbers of word candidates 6, 20, and 30 corresponding to the SN ratios of inputted speech 50dB, 30dB, and 10dB respectively are connected by a straight line, as shown in Figure 12, indicating an approximate value of each point and interpolating values in each SN ratio range by a straight linear function.

In this case, the interpolating equation deriver 25 uses the data in the number-of-word-candidates storing table 24, makes a linear regression analysis on the data of the value of an SN ratio and the corresponding necessary word candidates, interpolates values between predetermined points using a linear function, and gradually changes the number of word candidates such that each of the points indicating the necessary number of word candidates 6, 20, and 30 corresponding to the SN ratios 50dB, 30dB, and 10dB respectively is connected to adjacent points by a straight line, after adjusting each position a little to make a straight line graph. That is, in the third embodiment, the interpolating equation derived by the interpolating equation deriver 25 is represented as follows.

$$y = ax + b \quad (3)$$

where a and b are coefficients.

A linear equation approximation unit 252 obtains a and b, the coefficients in the linear equation y=ax+b, using the values of SN ratios and the number of candidates read from the number-of-recognition candidates storing table 24, that is, the relationship between the thresholds 50dB, 30dB, and 10dB and the corresponding numbers of recognition candidates 6, 20, and 30 according to the linear regression analysis based on the conventional method of least squares. If the coefficients a and b are obtained, the corresponding SN ratio of x_1 is substituted for x. Thus, as in the above described second embodiment, the number-of-recognition candidates determiner 26 determines the necessary number of recognition candidates y_1. In this case, $y_1 \fallingdotseq 25$ if $x_1 = 20$dB.

If the value of an SN ratio of inputted speech calculated by the SN ratio calculator 15 is 20dB, for example, the necessary number of word candidates is determined to be about 25 as in the above described second embodiment, and the recognized word candidate determiner 14 outputs 25 highest order word candidates as a recognition candidate word string. If the 25 highest order word candidates are prepared, the recognition rate of 90% can be guaranteed.

As described above, in the third embodiment, the total area of the value of an SN ratio outputted by the SN ratio calculator 15 is interpolated by a straight line represented by y = ax + b. Therefore, the necessary number of arithmetic operations can be considerably reduced. Besides, as in the above described second embodiment, the necessary number of word candidates is varied in proportion to the change in the value of an SN ratio of an inputted speech. Therefore, the necessary number of word candidates can be obtained corresponding to the obtained value of the SN ratio, thereby improving the efficiency of a recognizing operation.

Although the linear equation y = ax + b is used as an example in the third embodiment, the equation can be replaced with a high degree function such as $y = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + b$ to obtain coefficients $a_n, a_{n-1}, \cdots, a_1, b$ according to the multiple regression analysis based on the above described method of least squares. (However in this case, the value of n is limited by the range of data stored in the number-of-recognition candidates storing table 24.)

Although a word is recognized in each of the above described embodiments, other objects such as syllable, phrase and sentence can be processed as objects to be recognized.

When a phrase or a sentence (during conversation) is considered as an application of the objects to be recognized, the following sentence, for example, is to be recognized.

"Watashiwa Gakkoe Ikimasu"

(I go to school)

When the recognition is performed in units of phrase, the following candidates are considered for respective phrases.

1. Watashiwo Gakkowa Ikimasu
2. Watashini Gakkoe Itta
3. Watashiwa Gakkono Ikimashita

As shown above, the candidates are provided for respective phrases. This is performed based on a completely vocal process. Therefore, the phrase recognition or the sentence recognition can pick up a candidate which is always not ranked at a highest position, by using a syntax or semantics analysis. If a certain phrase is excluded from a candidate, it is very difficult to restore the excluded candidate at the stage of performing the syntax or semantics analysis.

On the other hand, as described above, the right candidate tends to be excluded from a group of candidates when the SN ratio decreases, and if the number of candidates are constant. Therefore, the present invention is advantageous in that the right candidate is prevented from being excluded from a group of the candidates by changing the number of the candidates based on the SN ratio.

As described above, the present invention removes the problem that a word to be reasonably selected as a recognition candidate may be excluded from the candidates due to an undesirable change in an SN ratio caused by an environmental change by changing the necessary number of word candidates in proportion to the value of an SN ratio, that is, a ratio of inputted speech to the environmental noise to guar-

BNSDOCID: <EP_____0526347A2_I_>

antee a predetermined recognition rate. Thus, words can be recognized at a high speed, thereby improving the efficiency in a recognizing operation.

Claims

1. A number-of-recognition candidates determining system in a speech recognition device for calculating the similarity between a plurality of preliminarily registered speeches and an inputted speech and outputting as recognition candidates a predetermined number of said registered speeches at higher order in similarity, comprising:

   an SN ratio calculator means (15) for calculating an SN ratio of inputted speech to the environmental noise, and

   a number-of-recognition candidates controller means (16) for controlling the number of said recognition candidates according to the value of said SN ratio calculated by said SN ratio calculator means.

2. The number-of-recognition candidates determining system according to Claim 1, wherein

   said number-of-recognition candidates controller means (16) comprises:

   an interpolating equation deriver means (25) for selecting a number of SN ratios each having an appropriately corresponding number of recognition candidates to obtain a predetermined recognition rate and for deriving an interpolating equation for interpolating the numbers of recognition candidates between said selected SN ratios, and

   a number-of-recognition candidates storing table means (24) for storing said number of recognition candidates corresponding to each of said SN ratios.

3. A number-of-recognition candidates determining system in a speech recognition device for calculating the similarity between a plurality of preliminarily registered speeches and outputting as recognition candidates a predetermined number of said registered speeches at higher order in similarity, comprising:

   an SN ratio calculator means (15) for calculating an SN ratio of an inputted speech to the environmental noise,

   an interpolating equation deriver means (25) for selecting SN ratios each having an appropriately corresponding number of recognition candidates to obtain a predetermined recognition rate and for deriving an interpolating equation for interpolating the numbers of recognition candidates between said selected SN ratios,

   a number-of-recognition candidates storing table means (24) for storing said number of recognition candidates corresponding to each of said SN ratios, and

   means for determining a number-of-recognition candidates by interpolating the necessary number of recognition candidates such that it is fixed as a constant value for an SN ratio between said selected SN ratios to determine said necessary number of recognition candidates corresponding to said SN ratio.

4. The number-of-recognition candidates determining system according to Claim 3, wherein

   said necessary number of recognition candidates for an SN ratio between said selected SN ratios is fixed to the necessary number of recognition candidates for the smaller value of said selected SN ratios.

5. A number-of-recognition candidates determining system in a speech recognition device for calculating the similarity between a plurality of speech preliminarily registered speeches and outputting as recognition candidates a predetermined number of said registered speeches at higher order in similarity, comprising:

   an SN ratio calculator means (15) for calculating an SN ratio of an inputted speech to the environmental noise,

   an interpolating equation deriver means (25) for selecting a number of SN ratios each having an appropriately corresponding number of recognition candidates to obtain a predetermined recognition rate and for deriving an interpolating equation for interpolating the numbers of recognition candidates between said selected points,

   a number-of-recognition candidates storing table means (24) for storing said number of recognition candidates corresponding to each of said SN ratios, and

   a number-of-recognition candidates determiner means (26) for determining said necessary number of recognition candidates by substituting a value of any SN ratio in said interpolating equation, wherein

   an interpolating equation is obtained such that said necessary number of recognition candidates is gradually varied in each SN ratio range between said selected SN ratios and that said interpolating equation determines said necessary number of recognition candidates for any SN ratio.

6. The number-of-recognition candidates determining system in a speech recognition device according to Claim 5, wherein

   said interpolating equation generates a linear change in a graph indicating the change in

BNSDOCID: <EP_____0526347A2_I_>

said necessary number of recognition candidates for each SN ratio between said selected SN ratios.

7. A number-of-recognition candidates determining system in a speech recognition device for calculating the similarity between a plurality of speech preliminarily registered speeches and outputting as recognition candidates a predetermined number of said registered speeches at higher order in similarity, comprising:

an SN ratio calculator means (15) for calculating a ratio of an inputted speech to the environmental noise (hereinafter referred to as an SN ratio),

an interpolating equation deliver means (25) for selecting a number of SN ratios each having an appropriately corresponding number of recognition candidates to obtain a predetermined recognition rate and for deriving an interpolating equation for interpolating the numbers of recognition candidates between said selected points,

a number-of-recognition candidates storing table means (24) for storing said number of recognition candidates corresponding to each of said SN ratios, and

a number-of-recognition candidates determiner means (26) for determining said necessary number of recognition candidates by substituting a value of any SN ratio in said interpolating equation, wherein

said necessary number of recognition candidates is determined for any SN ratio by obtaining an interpolating equation for gradually varying the number of recognition candidates for each of said SN ratio ranges between said selected SN ratios.

8. The number-of-recognition candidates determining system In a speech recognition device according to Claim 7, wherein

said interpolating equation for gradually varying said necessary number of recognition candidates for each of said SN ratio ranges between said selected SN ratios is a linear function equation.

9. The number-of-recognition candidates determining system in a speech recognition device according to Claim 7, wherein

said interpolating equation for gradually varying said necessary number of recognition candidates for each of said SN ratio ranges between said selected SN ratios is a high degree function equation obtained by a multiple regression analysis.

BNSDOCID: <EP_____0526347A2_I_>

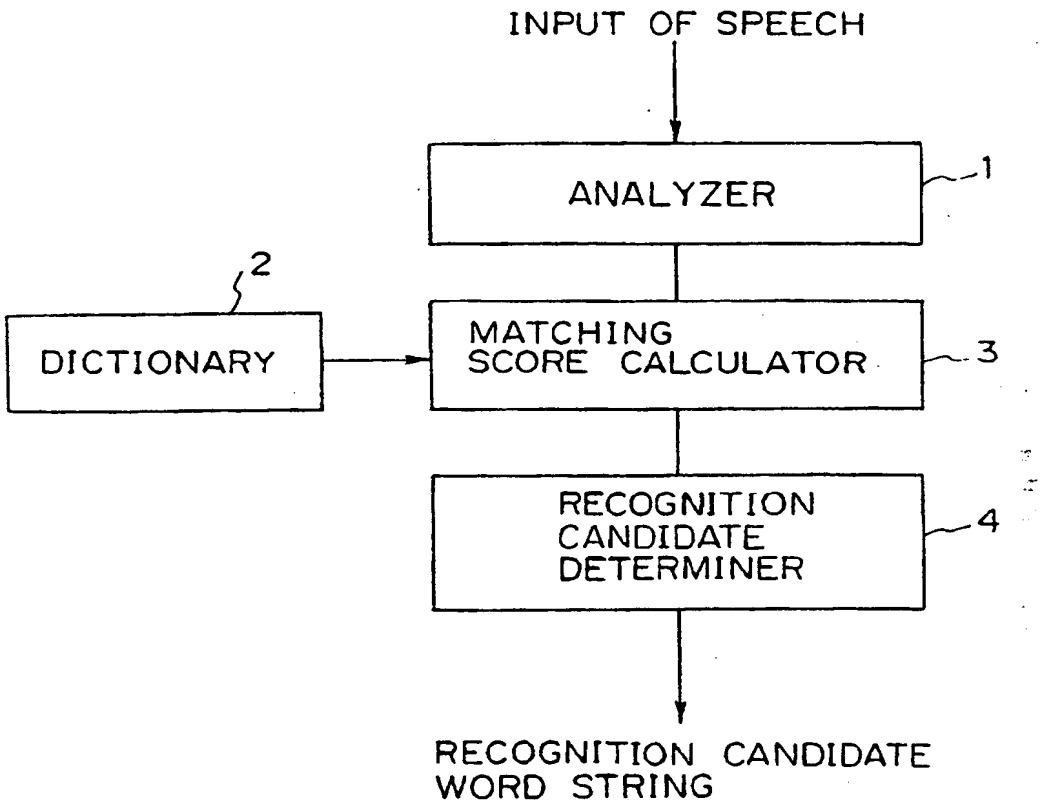

Fig. 1

BNSDOCID: <EP_____0526347A2_I_>

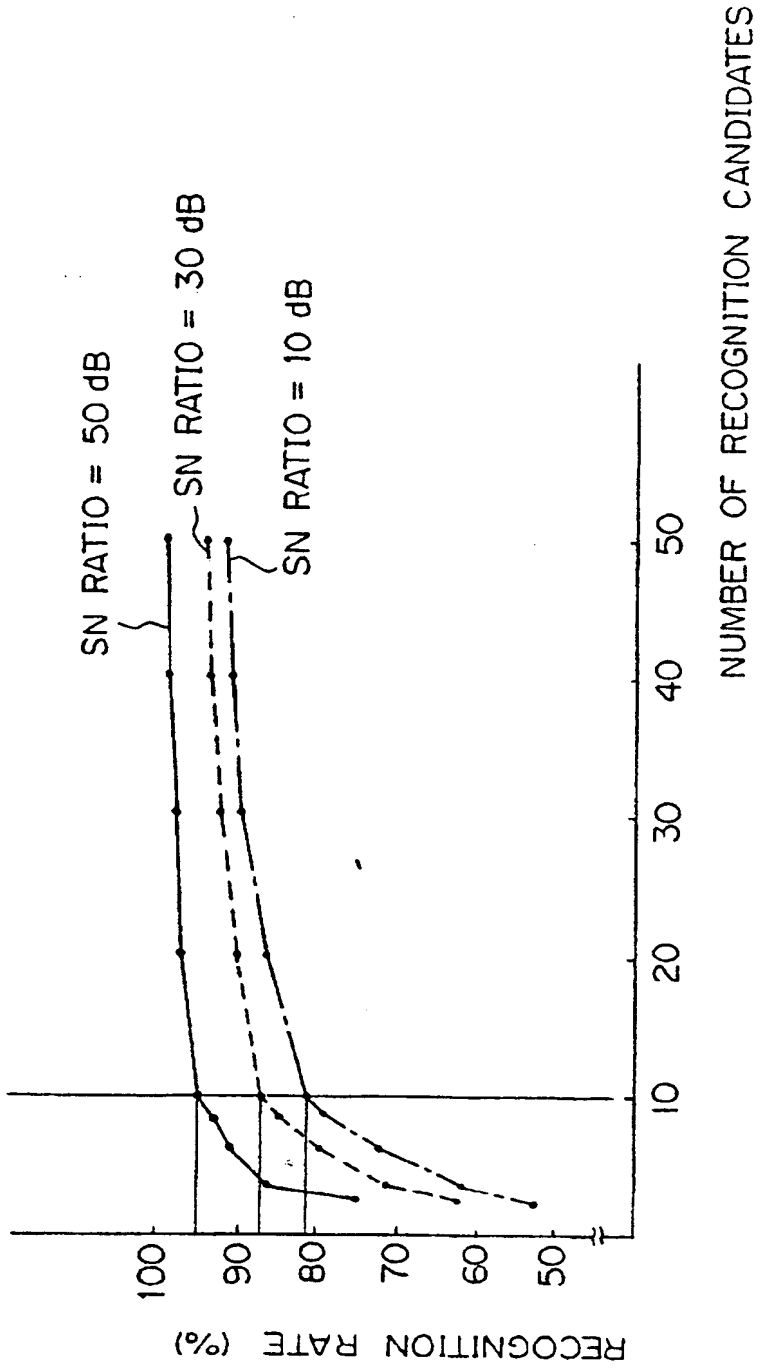

Fig. 2

BNSDOCID: <EP_____0526347A2_I_>

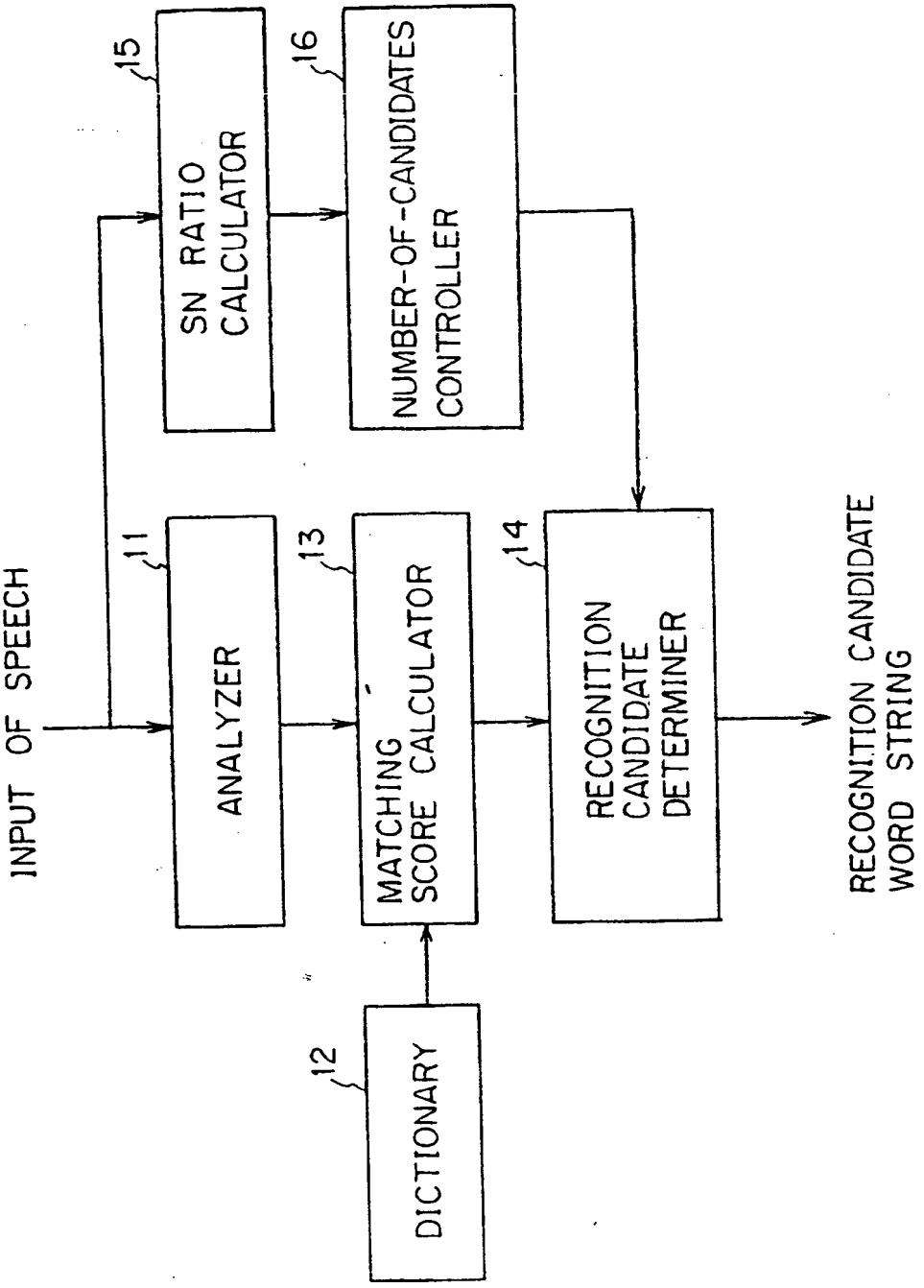

Fig. 3

BNSDOCID: <EP_____0526347A2_I_>

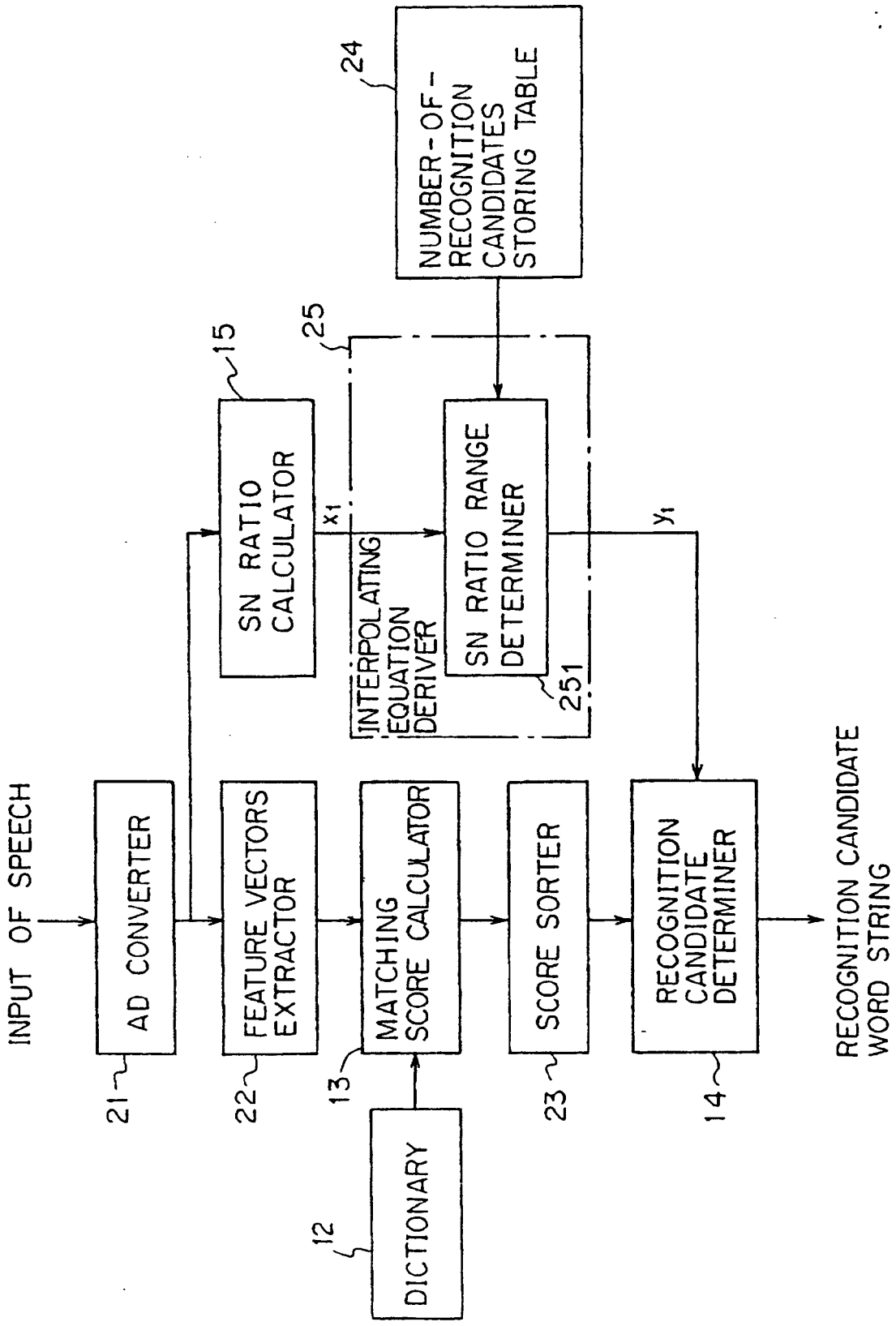

Fig. 4

BNSDOCID: <EP_____0526347A2_I_>

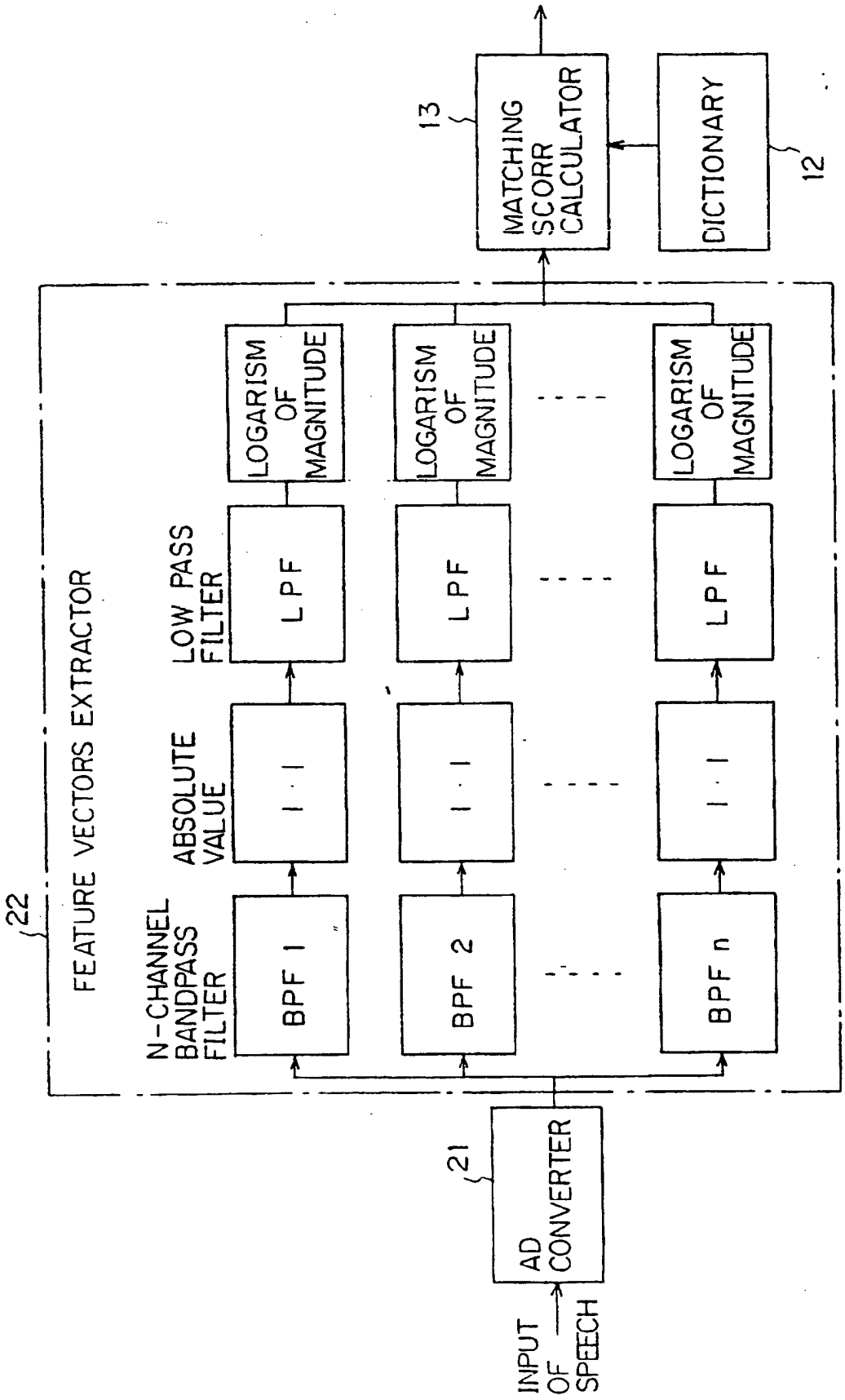

Fig. 5

BNSDOCID: <EP_____0526347A2_I_>

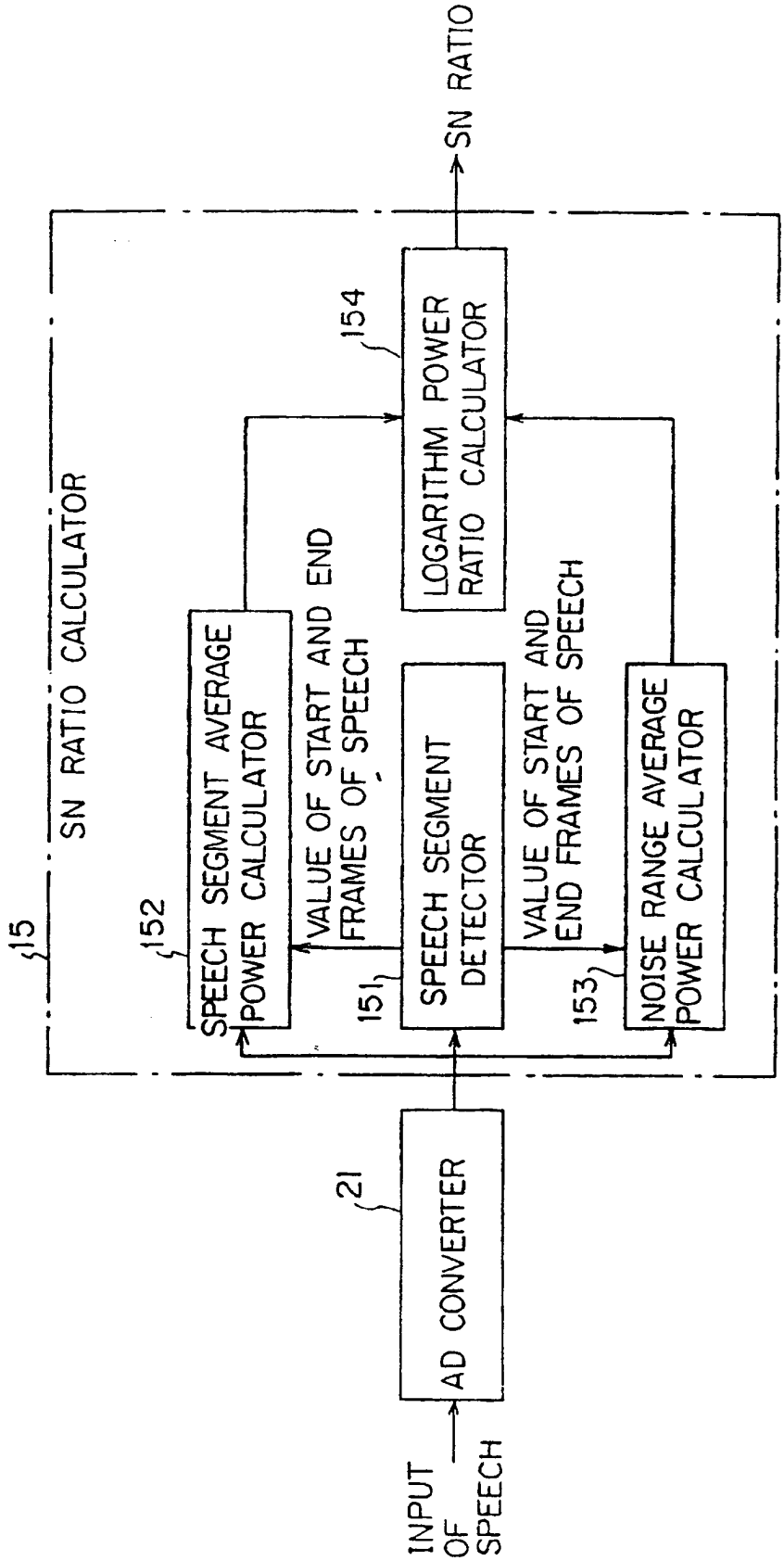

Fig. 6

BNSDOCID: <EP_____0526347A2_I_>

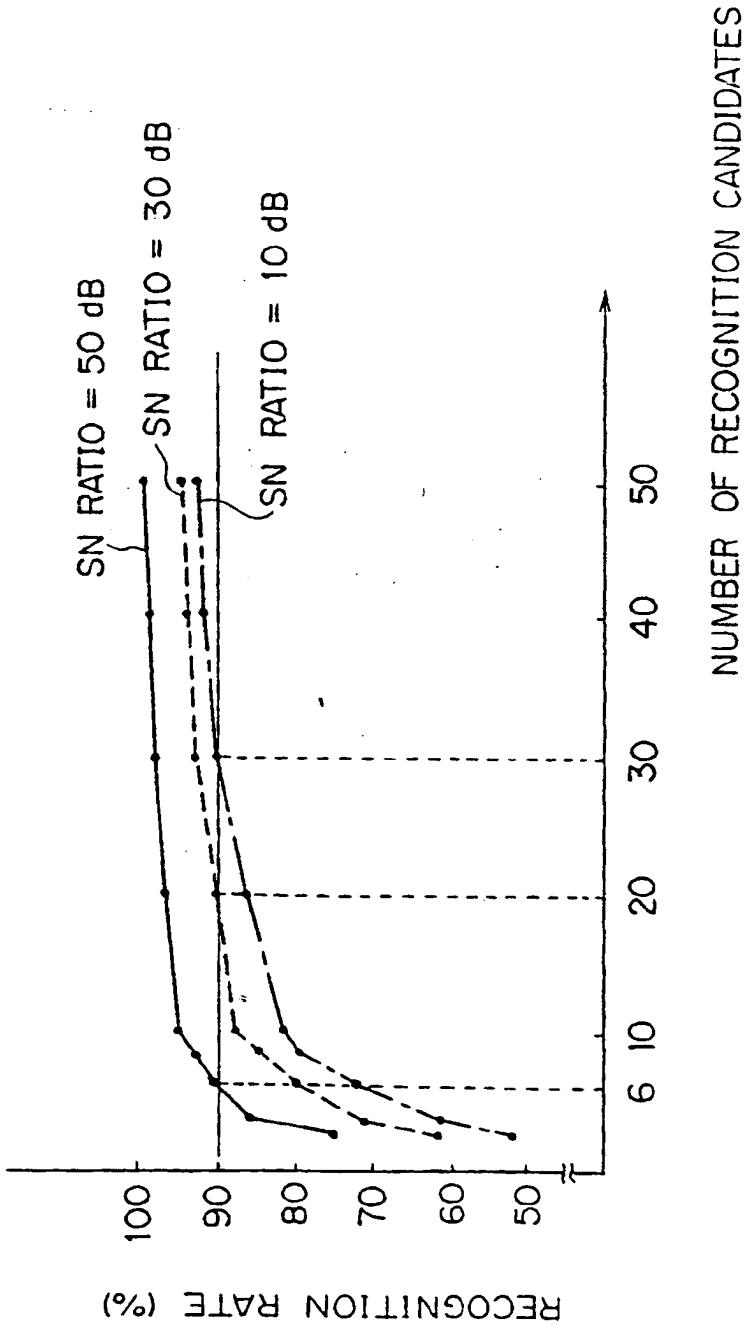

Fig. 7

BNSDOCID: <EP___0526347A2_I_>

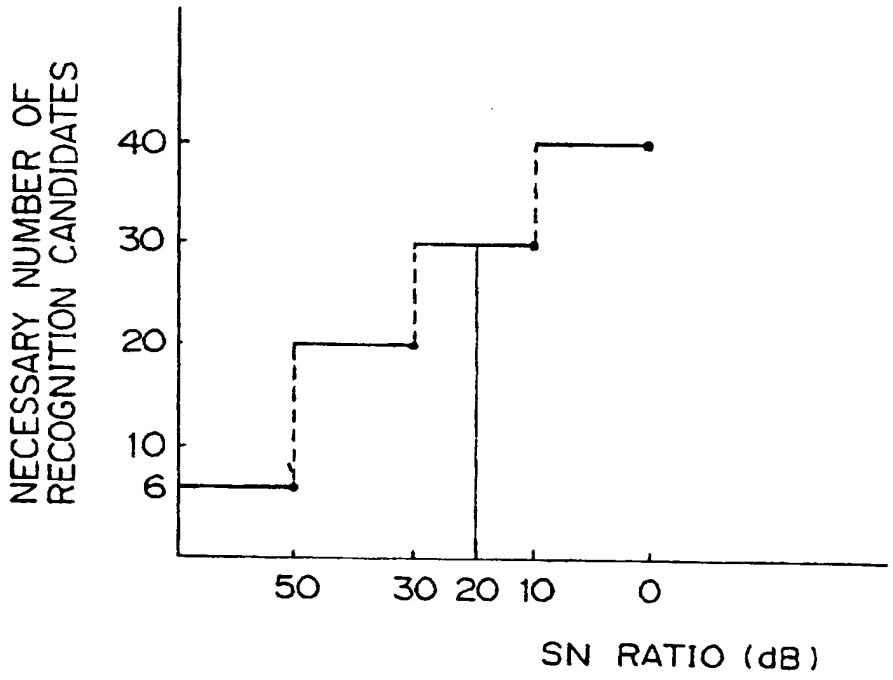

Fig. 8

BNSDOCID: <EP_____0526347A2_I_>

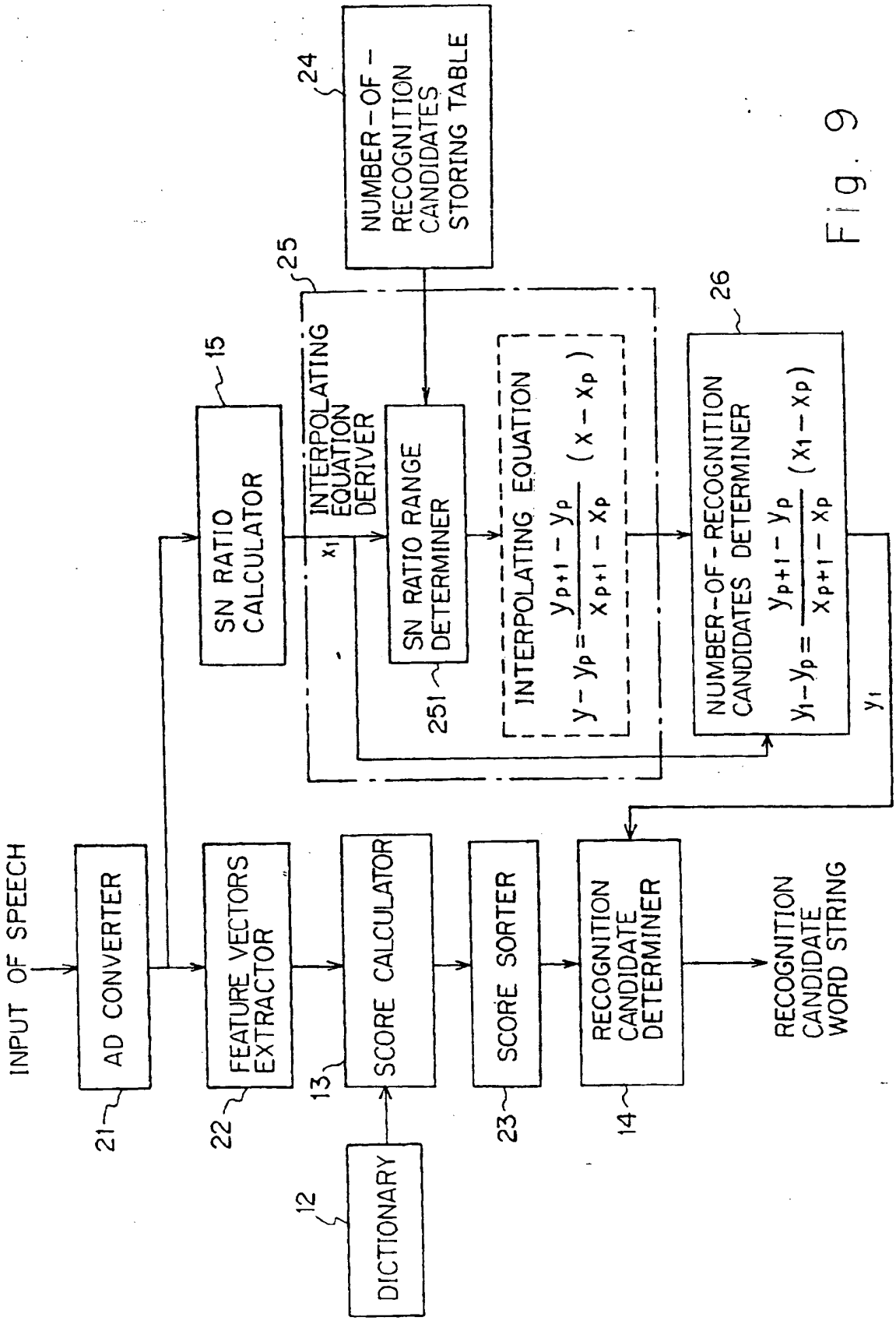

Fig. 9

BNSDOCID: <EP_____0526347A2_I_>

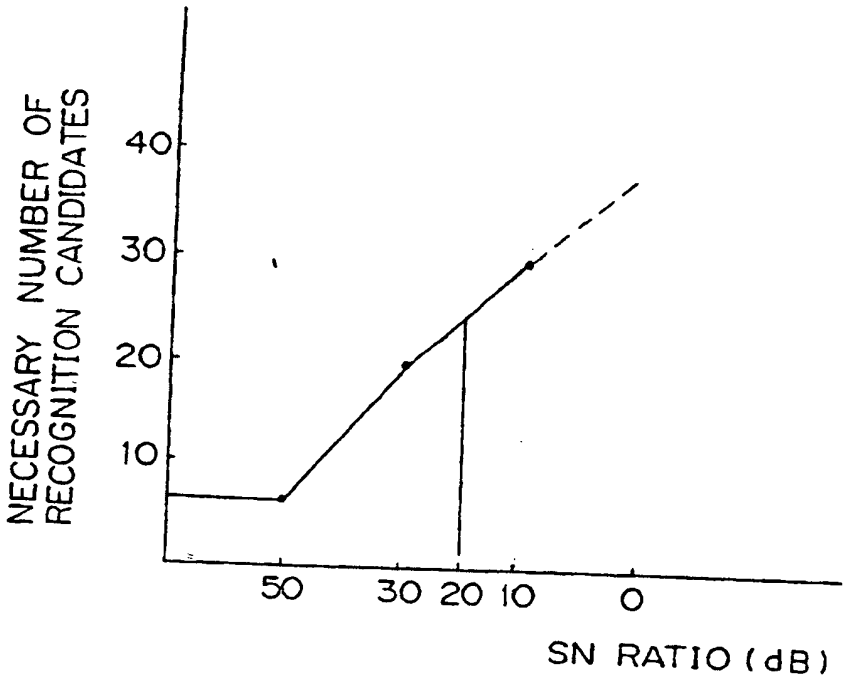

Fig. 10

BNSDOCID: <EP_____0526347A2_I_>

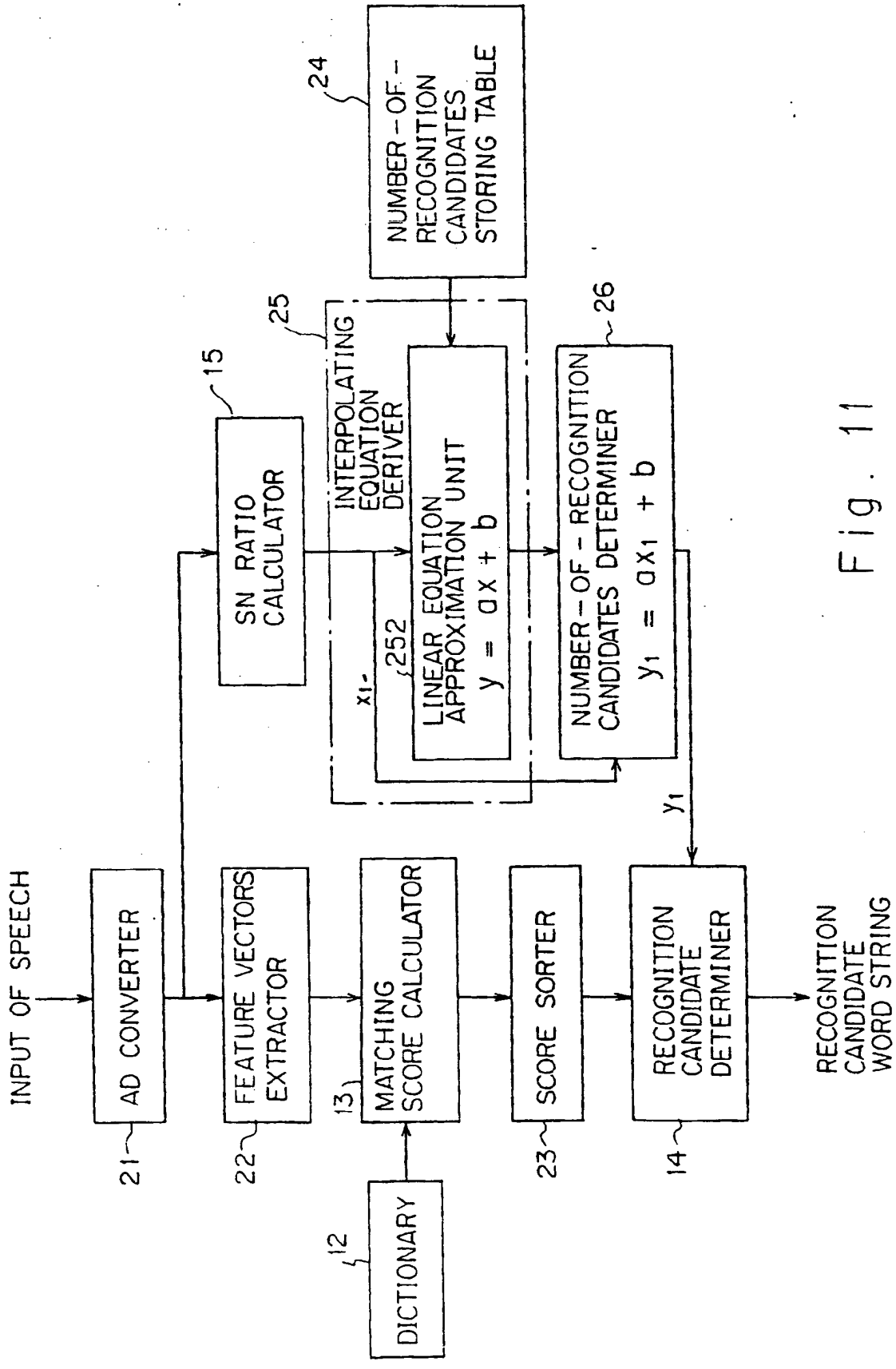

Fig. 11

BNSDOCID: <EP_____0526347A2_I_>

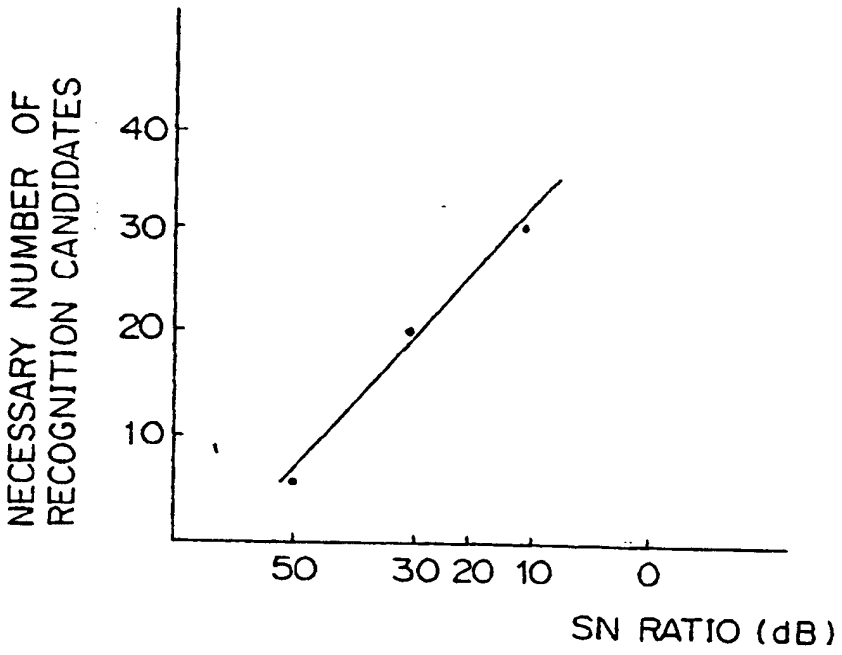

Fig. 12

BNSDOCID: <EP_-_____0526347A2_I_>